SIAM J. APPL. MATH.
Vol. 78, No. 5, pp. 2626–2647
© 2018 Society for Industrial and Applied Mathematics

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

STABILITY OF INVERSE TRANSPORT EQUATION IN DIFFUSION SCALING AND FOKKER–PLANCK LIMIT*

KE CHEN†, QIN LI‡, AND LI WANG§

Abstract. We consider the radiative transfer equation (RTE) with two scalings in this paper: one is the diffusive scaling whose macroscopic limit is a diffusion equation, and the other is a highly forward peaked scaling, wherein the scattering term is approximated by a Fokker–Planck operator as a limit. In the inverse setting, we are concerned with reconstructing the scattering and absorption coefficients using boundary measurements. As the measurement is often polluted by errors, both experimental and computational, an important question is to quantify how the error is amplified or suppressed in the process of reconstruction. Since the solution to the forward RTE behaves differently in different regimes, it is expected that stability of the inverse problem will vary accordingly. Particularly, we adopted the linearized approach and showed, in the former case, that the stability degrades when the limit is taken, following a similar approach as in [K. Chen, Q. Li and L. Wang, *Inverse Problems,* 34 (2018), 025004]. In the latter case, we showed that a full recovery of the scattering coefficient is less possible in the limit.

Key words. transport equation, diffusion limit, Fokker–Planck limit, stability, inverse problem

AMS subject classifications. 82B40, 65N21, 35L99

DOI. 10.1137/17M1157969

1. Introduction.

The radiative transfer equation (RTE) describes the dynamics of photon particles propagating in scattering and absorbing media [13]. A typical form reads as

$$\partial_t f + v \cdot \nabla_x f = \int_{\mathbb{S}^{d-1}} k(x, v, v') f(t, x, v') \mathrm{d}v' - \sigma(x, v) f(t, x, v)\,, \tag{1.1}$$

equipped with initial condition

$$f(0, x, v) = f^I(x, v) \tag{1.2}$$

and boundary condition

$$f|_{\Gamma_-} = \phi(t, x, v)\,. \tag{1.3}$$

Here $f(t, x, v)$ is the distribution of particles at location $x \in \Omega \subset \mathbb{R}^d$ moving with velocity $v \in \mathbb{S}^{d-1}$. Since photons travel with a fixed speed, the velocity v is normalized to $|v| = 1$. $k(x, v, v')$ is the scattering cross section, representing the probability of particles that move in direction v' changing to direction v. σ is the total scattering

*Received by the editors November 21, 2017; accepted for publication (in revised form) August 9, 2018; published electronically October 9, 2018.

http://www.siam.org/journals/siap/78-5/M115796.html

Funding: The first and the second authors are supported in part by a start-up fund from UW-Madison, NSF-DMS 1619778 and 1750488, TRIPODS 1740707, and UW-data science initiatives. The third author is supported in part by a start-up fund from SUNY-Buffalo, University of Minnesota, and NSF-DMS 1620135.

†Mathematics Department, University of Wisconsin–Madison, Madison, WI 53705 (ke@math.wisc.edu).

‡Mathematics Department and Wisconsin Institute of Discovery, University of Wisconsin–Madison, Madison, WI 53705 (qinli@math.wisc.edu).

§School of Mathematics, University of Minnesota, Twin Cities, Minneapolis, MN 55455 (wang8818@umn.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

coefficient that consists of the amount of photon particles being scattering and absorbed by the material. k and σ constitute the main optical property of the material.

The boundary condition (1.3) is a common choice for RTE, and Γ_- represents the "incoming" portion of the boundary ($\partial\Omega \times \mathbb{S}^{d-1}$), i.e.,

$$\Gamma_- = \{(x,v) : x \in \partial\Omega\,,\ v \in \mathbb{S}^{d-1}\,,\ v \cdot n_x < 0\}\,, \tag{1.4}$$

where n_x is the unit outer normal direction of the boundary. Similarly, one can define the "outgoing" portion of the boundary by

$$\Gamma_+ = \{(x,v) : x \in \partial\Omega\,,\ v \in \mathbb{S}^{d-1}\,,\ v \cdot n_x > 0\}\,. \tag{1.5}$$

The well-posedness of the forward problem (1.1)–(1.3) is summarized in [18].

RTE (1.1) is often incorporated with different scales that lead it to different equations. One typical scaling is the diffusive scaling, under which the RTE is well approximated by a diffusion equation

$$\partial_t \rho = C\nabla_x \left(\frac{1}{\sigma_s}\nabla_x \rho\right) + \sigma_a \rho\,,$$

where $\rho(t,x) = \int_{\mathbb{S}^{d-1}} f \mathrm{d}v$, and σ_s, σ_a (related to k and σ) will be defined later. C is a generic constant depending on the dimension of the problem. This scaling is encountered in the long time limit with a strong scattering effect. Another is the Fokker–Planck scaling which emphasizes the highly forward peaked scattering. In this case, (1.1) reduces to

$$\partial_t f + v \cdot \nabla_x f = \mathcal{L}_{\mathrm{FP}} f\,,$$

where

$$\mathcal{L}_{\mathrm{FP}} = \left[\frac{\partial}{\partial v_3}(1-v_3^2)\frac{\partial}{\partial v_3} + \frac{1}{1-v_3^2}\frac{\partial^2}{\partial \psi^2}\right]\,,$$

and $v = (\sqrt{1-v_3^2}\cos\psi, \sqrt{1-v_3^2}\sin\psi, v_3)$. In both scenarios, theory exists regarding the derivation, validity, and asymptotic error in the approximation, and the reader can refer to [23, 11] for the former case and [24, 28] for the latter.

We study the two limiting procedures of this problem in the inverse setting, with special attention paid to the stability of the reconstruction. Unlike the forward setting wherein the optical properties k and σ are given, and one attempts to solve $f(t,x,v)$ for a specific boundary condition (1.3), in the inverse problem, one tries to recover the unknown optical properties from boundary measurements of $f(t,x,v)$. To be more precise, we define the albedo operator as a mapping from the boundary condition $\phi(t,x,v)$ to the outgoing data $f|_{\Gamma_+}$:

$$\mathcal{A}(k,\sigma): \quad \phi \mapsto f|_{\Gamma_+}\,;$$

then by adjusting the incoming data ϕ, and measuring the corresponding outgoing data $f|_{\Gamma_+}$, one gains a full knowledge of $\mathcal{A}$, which can be used to determine k and σ.

The inverse RTE problem benefits a broad application in optical tomography, atmospheric science, and aerospace engineering. Optical tomography, with its major application in medical imaging, utilizes scattered light as a probe of structural variations in the optical properties of the tissue. Specifically, a narrow collimated beam of low energy visible or near infrared light is sent into biological tissues, then collected by

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

an array of detectors after it propagates through the media. The measurements collected are used to recover the optical properties of the media. In atmospheric science, or remote sensing, satellites cumulate hyperspectral light reflected from the earth, which is used to infer mineral or plant distribution on the ground. In aerospace engineering, pictures taken by spacecraft in the universe (Galileo's pictures from Jupiter, or Cassini's pictures from Saturn, for example) are sent back to Earth for analyzing mineral/gas distribution on different planets. In all applications, the forward solver for light propagation is described by the RTE. One measures the reflected or propagated light intensity to reconstruct the optical properties, with which tissue/ground/gas components are inferred.

On the analytical side, there has been a vast literature on the well-posedness and stability of the inverse problem. In a pioneering paper [17], the authors showed that both k and σ can be uniquely determined by the incoming-to-outgoing map $\mathcal{A}$, assuming that σ is v-independent. With v dependence, the uniqueness up to the gauge-invarience was shown in [33]. The analysis is done through performing the singular decomposition: one separates the collected $f|_{\Gamma_+}$ data according to the singularities, and different parts are in charge of recovering different coefficients. Another approach is to linearize the equation before applying inverse Born series, followed by showing the convergence of the series [26]. The results on the stability of the "inverse" dates back to [35] and was made systematic in [7, 8, 9]. Many papers concern the time-dependent case and the associated stability analysis has also been conducted [22, 31, 16, 6, 10]; see also [5] for a review.

On the numerical side, special care is needed to address the ill-posedness of the problems, both inherited from the continuous counterpart and due to the incomplete corrupted data. Indeed, to uniquely determine k and σ, one needs a full knowledge of $\mathcal{A}$, which is impractical in real applications, and measurement error can easily propagate and get exaggerated. Typically Tikhonov type regularization is used to balance the pollution and the error tolerance, and the type of regularization embeds some prior knowledge. See [1] using the standard L^2, [34] using total variation regularization for the least variance, [29] summarizing H^1 regularization for some regularity, and L^1 regularization for sparsity. See also Tikhonov type regularization used on each element in the inverse Born series [25, 26]. Besides the ill-posedness, the size of the problem also brings extra difficulties, and Jacobian type techniques [32] are introduced to advance the computation.

In the presence of different scales, however, the above mentioned theory or algorithms cannot be directly applied since the inverse problem may completely change its type. One example is the diffusive regime; while the inverse RTE with sufficient variation in measurement is shown to be well-posed, its diffusion limit is the Calderón type problem which is well acknowledged to be ill-posed [19, 20]. Our goal in this paper is to provide a rigorous connection between different scalings in the inverse setting and show how stability varies with the scaling parameter. For the diffusive scaling, the connection is observed in [30, 2, 3] and addressed in [4, 14]. A similar problem on recovering the doping profile in the Boltzmann–Poisson system is presented [15], wherein numerical simulations also imply this relation. In the Fokker–Planck regime, the limit was briefly mentioned in [5] but the full discussion was rarely seen in the literature. In this paper, we restrict our attention to the linearized version of the inverse problem and mainly contribute in the following two aspects:

(1) extend our previous analysis with diffusive scaling for steady problems [14] to time-dependent problems, and show that the stability degrades in the diffusion limit;

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(2) examine the well-posedness and stability in the Fokker–Planck scaling which has never been studied in detail before.

In the former case, although the conclusion here looks similar to that in the steady state scenario [14], the extra dependence on time calls for variations in technical details, as evident from the construction of adjoint problem (e.g., (2.13)), possible presence of initial layer (e.g., Remark 2.2), and property of the limiting solution (e.g., proof of Theorem 2.5). In the latter case, we hereby set up the inverse problem in detail and analyze the stability rigorously for the first time.

The rest of paper is organized as follows. Section 2 is devoted to the diffusion regime and section 3 is devoted to the forward peaked regime, in which the diffusion equation and the Fokker–Planck equation are obtained as asymptotic limit, respectively. In both cases, we utilize the linearization approach, study the well-posedness, and examine the change of stability while passing the limit.

2. Diffusion regime. In this section, we first briefly recapitulate the properties of RTE and its diffusion limit, and then examine the well-posedness in the inverse setting. For simplicity, we assume that the optical properties only have spatial dependence, and we rewrite (1.1) as

$$\partial_t f + v\cdot\nabla_x f = \frac{\sigma_s(x)}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{d-1}}(f(t,x,v')-f(t,x,v))\mathrm{d}v' - \sigma_a(x)f(t,x,v)\,, \tag{2.1}$$

where σ_s is termed the scattering coefficient, and

$$\sigma_a(x) = \sigma(x) - \sigma_s(x)$$

is the absorption coefficient. Here $|\mathbb{S}^{d-1}|$ is the area of unit sphere in dimension $d-1$.

The diffusion regime is achieved in the long time limit and when the scattering is much stronger than the absorption. Indeed, we introduce a small parameter—Knudsen number Kn—and rescale time and space as

$$t \to \frac{t}{\mathsf{Kn}^2}, \quad x \to \frac{x}{\mathsf{Kn}}\,.$$

We also let $\sigma_a \to \mathsf{Kn}\sigma_a$ so that the absorption effect is negligible compared to the scattering effect; then the RTE is rewritten as

$$\begin{cases} \mathsf{Kn}\partial_t f + v\cdot\nabla_x f = \frac{1}{\mathsf{Kn}}\sigma_s\mathcal{L}f - \mathsf{Kn}\sigma_a f \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,,\\ f(0,x,v) = f^I(x,v)\,,\\ f|_{\Gamma_-}(t,x,v) = \phi(t,x,v)\,. \end{cases} \tag{2.2}$$

Here the collision operator $\mathcal{L}$ is an abbreviation of

$$\mathcal{L}f(t,x,v) = \frac{1}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{d-1}}(f(t,x,v')-f(t,x,v))\mathrm{d}v' = \langle f\rangle_v - f\,. \tag{2.3}$$

Note that the notation $\langle\cdot\rangle_v$ means taking the average in the velocity domain. There are two key features of the collision operator:

- Mass conservation: $\int \mathcal{L}f\mathrm{d}v = 0$.
- One-dimensional null space: By setting $\mathcal{L}f = 0$, one gets $f = \frac{1}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{d-1}} f\mathrm{d}v$, meaning that f is a constant in velocity domain. We denote it as $\mathrm{Null}\,\mathcal{L} = \mathrm{span}\{\rho(t,x)\}$, the collection of functions that depend on t and x only.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.1. Diffusion limit. When $\mathsf{Kn} \ll 1$, the equation falls into the diffusion regime, and the RTE is asymptotically equivalent to a diffusion equation as $\mathsf{Kn} \to 0$.

THEOREM 2.1 (see [23, 27, 12]). *Suppose f solves (2.2) with initial data $f(0,x,v) = f^I(x)$ and boundary data $f|_{\Gamma_-} = \phi(t,x)$, both of which are independent of velocity v. Then as $\mathsf{Kn} \to 0$, $f(t,x,v)$ converges to $\rho(t,x)$, which solves the heat equation:*

$$\begin{cases} \partial_t \rho - C\nabla_x \cdot \left(\frac{1}{\sigma_s}\nabla_x \rho\right) + \sigma_a \rho = 0\,, \\ \rho(0,x) = f^I(x)\,, \\ \rho|_{\partial\Omega} = \phi(t,x)\,. \end{cases} \tag{2.4}$$

Here C is a time-dependent constant.

Here we omit the rigorous proof but only provide the following asymptotic analysis as it is the major building block. In the zero limit of Kn, the distribution converges to the local equilibrium, and by applying the standard asymptotic expansion technique, we write

$$f_{\text{in}} = f_0 + \mathsf{Kn} f_1 + \mathsf{Kn}^2 f_2 + \cdots\,. \tag{2.5}$$

Insert the expansion in (2.2) and equate like powers of Kn:

$\mathcal{O}(1)$ $\mathcal{L}f_0 = 0$. This immediately indicates that $f_0 \in \text{Null}\,\mathcal{L}$. With the form given in (2.3), $\text{Null}\,\mathcal{L}$ consists of functions that are constants in the v domain, and thus $f_0(t,x,v) = \rho(t,x)$.

$\mathcal{O}(\mathsf{Kn})$ $v \cdot \nabla_x f_0 = \sigma_s \mathcal{L} f_1$. This indicates that $f_1 = \mathcal{L}^{-1}\frac{v\cdot\nabla_x f_0}{\sigma_s}$. Notice that $\mathcal{L}$ is invertible on $\text{Null}\,\mathcal{L}^\perp$, and consider the form of $\mathcal{L}$ in (2.3); then $\text{Null}\,\mathcal{L}^\perp = \{f : \int f \mathrm{d}v = 0\}$ and $\frac{v\cdot\nabla_x f_0}{\sigma_s} \in \text{Null}\,\mathcal{L}^\perp$, and therefore $f_1 = \mathcal{L}^{-1}\frac{v\cdot\nabla_x f_0}{\sigma_s} = -\frac{v\cdot\nabla_x f_0}{\sigma_s}$.

$\mathcal{O}(\mathsf{Kn}^2)$ $\partial_t f_0 + v\cdot\nabla_x f_1 = \sigma_s \mathcal{L} f_2 - \sigma_a f_0$. Here we integrate the equation with respect to v. The second corrector f_2 will vanish and the left-hand side (LHS) becomes

$$\begin{aligned} &\partial_t \rho + \frac{1}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{d-1}} v\cdot\nabla_x\left(-\frac{v}{\sigma_s}\cdot\nabla_x\rho\right)\mathrm{d}v = -\sigma_a\rho \\ &\qquad\Rightarrow\quad \partial_t\rho - C\nabla_x\cdot\left(\frac{1}{\sigma_s}\nabla_x\rho\right) = -\sigma_a\rho\,. \end{aligned} \tag{2.6}$$

Integrating $\int v\cdot v\mathrm{d}v$ out, we obtain the diffusion limit and conclude the theorem. The constant C depends on the dimension of the velocity space.

Remark 2.2. We comment that the initial and boundary conditions in the above theorem are rather strict—they both are independent of v—so that no initial or boundary layers will be generated and ρ takes the same initial and boundary conditions as f. For more general cases, however, one needs to introduce the initial and boundary layers to damp out the nonhomogeneities. More specifically, we write

$$f(t,x,v) = f_{il}(t,x,v) + f_{bl}(t,x,v) + f_{int}(t,x,v)\,, \tag{2.7}$$

where $f_{int}(t,x,v)$ stands for the interior solution (i.e., x is away from the boundary and t is away from 0), and we write

$$f_{int}(t,x,v) = \theta(t,x) - \mathsf{Kn}\mathcal{L}^{-1}(v)\partial_x\theta(t,x) + \mathcal{O}(\mathsf{Kn}^2)\,,$$

with θ satisfying the diffusion equation (2.4) (but with different initial and boundary conditions than in (2.4)). f_{il} is the initial layer and is governed by

$$\partial_\tau f_{il} - \mathcal{L} f_{il} = 0\,, \tag{2.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $\tau = t/\mathsf{Kn}^2$ is the rescaled time. With appropriate initial data, f_{il} damps to 0 exponentially fast in τ and thus $f_{il} \sim e^{-t/\mathsf{Kn}^2} \sim 0$ for finite t with $\mathsf{Kn} \to 0$. f_{bl} is the boundary layer. At each point on the boundary, $x_0 \in \partial\Omega$, f_{bl} satisfies

$$v\partial_z f_{bl} + \mathcal{L} f_{bl} = 0\,, \tag{2.9}$$

where z is the rescaled spatial coordinate around x_0: $z = -\frac{(x-x_0)\cdot n_x}{\mathsf{Kn}}$ with n_x being the normal direction pointing out of Ω at x_0. It has been shown that f_{bl} exponentially decays to a constant in z. This constant is termed the extrapolation length and is uniquely determined by the boundary data around x_0. We denote it $\phi(t, x_0)$. This means that for x adjacent to x_0, in the zero limit of Kn, $|f_{bl} - \phi(t, x_0)| \sim e^{-|x-x_0|/\mathsf{Kn}} \sim 0$. We typically subtract this constant from f_{bl} and set it as the Dirichlet boundary condition for θ, and thus $f_{bl} \sim 0$ everywhere.

2.2. Recover absorption coefficient σ_a. In this section we assume that the scattering coefficient is known and aim to recover the absorption coefficient. Without loss of generality, we let $\sigma_s \equiv 1$.

2.2.1. Inverse problem setup. We first rewrite (2.2) into

$$\begin{cases} \mathsf{Kn}\partial_t f + v\cdot\nabla_x f = \frac{1}{\mathsf{Kn}}\mathcal{L} f - \mathsf{Kn}\sigma_a f \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{d-1}\,, \\ f(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{d-1}\,, \\ f(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,. \end{cases} \tag{2.10}$$

The solution to the above equation, denoted by $f(t,x,v;\phi)$, models the number density of photons with certain inflow ϕ. In experiment, time-dependent velocity-averaged data $m(t,x) := \int_{\mathbb{S}^{n-1}} v\cdot n(x) f(t,x,v)|_{\Gamma_+}\mathrm{d}v$ is collected on the out-flow boundary Γ_+. Therefore, we can define the albedo operator as

$$\mathcal{A}(\sigma_a): \quad \phi(t,x,v) \to m(t,x) = \int_{\mathbb{S}^{n-1}} v\cdot n(x) f(t,x,v)|_{\Gamma_+}\mathrm{d}v\,.$$

Notice that $\mathcal{A}$ nonlinearly depends on σ_a via the solution $f(t,x,v;\phi)$. In an effort to study the property of $\mathcal{A}$, we first derive a linearized version of it following the procedure outlined in [29]. For the angularly averaged measurement we considered here, see also [21] for the reconstruction of a stationary transport equation.

Suppose a priori information about the absorption coefficient is known in the sense that σ_a can be considered as a small perturbation around a background state $\sigma_{a0}(x)$, i.e.,

$$\sigma_a(x) = \sigma_{a0}(x) + \tilde{\sigma}_a(x) \quad \text{with} \quad |\tilde{\sigma}_a| \ll |\sigma_a|\,, \quad \text{a.e.}\,;$$

then a linearized problem with background state σ_{a0} and same initial and boundary data can be defined as

$$\begin{cases} \mathsf{Kn}\partial_t f_0 + v\cdot\nabla_x f_0 = \frac{1}{\mathsf{Kn}}\mathcal{L} f_0 - \mathsf{Kn}\sigma_{a0} f_0 \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{d-1}\,, \\ f_0(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{d-1}\,, \\ f_0(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,. \end{cases} \tag{2.11}$$

Comparing (2.10) and (2.11), we define the residue $\tilde{f} = f - f_0$; then it solves, to the leading order,

$$\begin{cases} \mathsf{Kn}\partial_t \tilde{f} + v\cdot\nabla_x \tilde{f} = \frac{1}{\mathsf{Kn}}\mathcal{L}\tilde{f} - \mathsf{Kn}\sigma_{a0}\tilde{f} - \mathsf{Kn}\tilde{\sigma}_a f_0 \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,, \\ \tilde{f}(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{n-1}\,, \\ \tilde{f}(t,x,v) = 0 \quad \text{on} \quad (0,T)\times\Gamma_-\,, \end{cases} \tag{2.12}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

which is obtained by subtracting (2.11) from (2.10) with the higher order term $\tilde{\sigma}_a \tilde{f}$ omitted. Notice here that both the linearized solution f_0 and the residue $\tilde{f}$ are implicitly dependent on the incoming data ϕ. We then introduced an adjoint problem of (2.11) and assign a Dirac delta function $\delta(\tau, y)$ at the boundary $(0,T)\times\Gamma_+$:

$$\begin{cases} -\mathsf{Kn}\partial_t g - v\cdot\nabla_x g = \frac{1}{\mathsf{Kn}}\mathcal{L}g - \mathsf{Kn}\sigma_{a0} g \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{d-1}\,, \\ g(T,x,v) = 0 \quad \text{on} \quad \{t=T\}\times\Omega\times\mathbb{S}^{d-1}\,, \\ g(t,x,v) = \delta(\tau,y) \quad \text{on} \quad (0,T)\times\Gamma_+\,. \end{cases} \tag{2.13}$$

The solution is denoted by $g(t,x,v;\tau,y)$. Multiplying (2.12) with g, (2.13) with $\tilde{f}$, integrating over $(0,T)\times\Omega\times\mathbb{S}^{n-1}$, and then subtracting them, we get

$$\int_{\Gamma_+(y)} \tilde{f}(\tau,y,v) n(y)\cdot v \mathrm{d}v = -\mathsf{Kn}\int_\Omega \tilde{\sigma}_a(x)\int_{\mathbb{S}^{n-1}}\int_0^T f_0(t,x,v;\phi) g(t,x,v;\tau,y)\mathrm{d}t\mathrm{d}v\mathrm{d}x\,, \tag{2.14}$$

where we have used the self-adjoint property of $\mathcal{L}$ and the divergence theorem $\int_{\mathbf{S}^{d-1}\times\Omega}\nabla_x\cdot(v\tilde{f}g)\mathrm{d}x\mathrm{d}v = \int_{\mathbb{S}^{d-1}\times\partial\Omega}\hat{n}\cdot v\tilde{f}g\ \mathrm{d}\Gamma$. We denote the LHS of (2.14) by $b(\tau,y,\phi)$; then according to the definition of $\tilde{f}$, it is simply

$$b(\tau,y,\phi) := \int_{\Gamma_+(y)} f(\tau,y,v) n(y)\cdot v\mathrm{d}v - \int_{\Gamma_+(y)} f_0(\tau,y,v) n(y)\cdot v\mathrm{d}v\,, \tag{2.15}$$

with the first term being the measurement from experiments and the second term computed from (2.11). This term therefore is known ahead of time. The RHS of (2.14) defines a linear mapping of $\tilde{\sigma}_a$. Let us denote

$$\gamma_{\mathsf{Kn}}(x;\tau,y,\phi) := -\mathsf{Kn}\int_{\mathbb{S}^{n-1}}\int_0^T f_0(t,x,v;\phi) g(t,x,v;\tau,y)\mathrm{d}t\mathrm{d}v; \tag{2.16}$$

then (2.14) defines a family of linear mapping from γ_{Kn} to the data on the LHS, parametrized by (τ,y,ϕ):

$$\int_\Omega \tilde{\sigma}_a(x)\gamma_{\mathsf{Kn}}(x;\tau,y,\phi)\mathrm{d}x = b(\tau,y,\phi)\,. \tag{2.17}$$

Therefore, (2.17) defines a linearized albedo operator, from which $\tilde{\sigma}_a$ can be obtained via solving a system of linear equations.

Remark 2.3. Equation (2.17) is a first type Fredholm operator, and it holds true for all parameter choices of τ, y, and ϕ. The study on the well-posedness simply relies on the space spanned by $\{\gamma_{\mathsf{Kn}}\}$. Suppose we look for $\tilde{\sigma}_a \in L_p(\mathrm{d}x)$; then the uniqueness is guaranteed if $\{\gamma_{\mathsf{Kn}}\}$ spans L_q space (with $\frac{1}{p}+\frac{1}{q}=1$). There have been many studies on the topic and this is not the main goal of the current paper. The well-posedness amounts to analyze the "conditioning" of γ_{Kn}. It is closely related to studying its "singular values," as will be explained in better detail below.

2.2.2. Ill-conditioning in the diffusion limit. Given the linearized albedo operator defined in (2.17), studying the stability of recovering $\tilde{\sigma}_a$ boils down to examining the property of the Fredholm operator of the first kind defined there. In this section, we intend to explore its conditioning with respect to Kn. More precisely, given a family of input-measurement pairs $(\phi(t,x,v),\ m(\tau,y,\phi))$, where $(t,x,v)\in(0,T)\times\Gamma_-$ and

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$m(\tau, y, \phi) = \int_{\mathbb{S}^{n-1}} f(\tau, y, v) n(y) \cdot v \mathrm{d}v$, we can explicitly compute $\gamma_{\mathsf{Kn}}(x)$ and $b(\tau, y, \phi)$ defined in (2.16) and (2.15) and study their dependence on Kn so as to get a sensitivity in recovering $\tilde{\sigma}_a$ with respect to Kn.

In this regard, we first introduce a distinguishability coefficient to quantify the perturbation of $\tilde{\sigma}_a$ when a δ-error is allowed for $b(\tau, y, \phi)$.

DEFINITION 2.4. *Consider linear equations* (2.17) *and* γ_{Kn} *defined in* (2.16) *and* $b(\tau, y, \phi)$ *defined in* (2.15)*; we define the distinguishability coefficient as*

$$\kappa_a := \sup_{\hat{\sigma}_a \in \Gamma_\delta} \frac{\|\hat{\sigma}_a - \tilde{\sigma}_a\|_{L^\infty(\mathrm{d}x)}}{\|\tilde{\sigma}_a\|_{L^\infty(\mathrm{d}x)}}, \tag{2.18}$$

where

$$\Gamma_\delta = \{\hat{\sigma}_a : \sup_{\substack{\forall \|\phi\|_{L^\infty(\Gamma_-)} \le 1, \\ \forall y \in \partial\Omega,\ \tau \in [0,T]}} |\langle \gamma_{\mathsf{Kn}}, \hat{\sigma}_a \rangle_{L^2(\mathrm{d}x)} - b(\tau, y, \phi)| \le \delta\},$$

and $\tilde{\sigma}_a$ *is the exact solution to* (2.17).

Here Γ_δ consists of all possible solutions to (2.17) within δ-tolerance, and the distinguishability coefficient κ quantifies supremum of relative error over Γ_δ. Therefore, in practice small κ is desired. However, this is not the case when Kn is small, as will be shown in the following theorem: small κ leads to very bad distinguishability.

THEOREM 2.5. *For a family of linear equations defined in* (2.17) *and an error tolerance* $\delta > 0$ *on the measurement, the distinguishability coefficient satisfies*

$$\kappa_a \ge \mathcal{O}\left(\frac{\delta}{\mathsf{Kn}}\right) \quad \textit{when} \quad \mathsf{Kn} \ll 1.$$

Proof. Let $c(x)$ be an arbitrary function that vanishes in the boundary layer and satisfies

$$\left| \int_\Omega \gamma_{\mathsf{Kn}}(x) c(x) \mathrm{d}x \right| \le \delta, \tag{2.19}$$

where γ_{Kn} is defined in (2.16). Choose $\hat{\sigma}_a = c(x) + \tilde{\sigma}_a$, then $\hat{\sigma}_a \in \Gamma_\delta$. When $\mathsf{Kn} \ll 1$, from Theorem 2.1, $f_0(t, x, v)$ can be decomposed into two parts,

$$f_0(t, x, v) = f_{\mathrm{int}}(t, x, v) + f_{\mathrm{bl}}(t, x, v),$$

where $f_{\mathrm{bl}}(t, x, v)$ encodes the boundary layer supported near the boundary with $\mathcal{O}(\mathsf{Kn})$ width and f_{int} is the interior solution, and it approaches its diffusion limit $\rho_f(t, x)$ which satisfies (2.4) with zero initial data and suitable boundary condition. Specifically, $f_{\mathrm{int}}(t, x, v)$ can be expanded as

$$f_{\mathrm{int}}(t, x, v) = \rho_f(t, x) - \mathsf{Kn} v \cdot \nabla_x \rho_f(t, x) + \mathcal{O}(\mathsf{Kn}^2), \tag{2.20}$$

where ρ_f solves

$$\begin{cases} \partial_t \rho_f = C \Delta_x \rho_f - \sigma_a \rho_f \quad \text{in} \quad (0, T) \times \Omega, \\ \rho_f(0, x) = 0 \quad \text{on} \quad \{t = 0\} \times \Omega, \\ \rho_f(t, x) = \eta_\phi(t, x) \quad \text{on} \quad (0, T) \times \partial\Omega. \end{cases} \tag{2.21}$$

Here the boundary value $\eta_\phi(x)$ is computed from $\phi(t, x, v)$ through the boundary layer analysis. (Details are provided in Remark 2.2.)

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Likewise, g admits the same decomposition that separates the interior part from the boundary part,

$$g_0(t,x,v) = g_{\text{int}}(t,x,v) + g_{\text{bl}}(t,x,v)\,,$$

and $g_{\text{int}}(t,x,v)$ has the expansion

$$g_{\text{int}}(t,x,v) = \rho_g(t,x) - \mathsf{Kn} v\cdot\nabla_x\rho_g(t,x) + \mathcal{O}(\mathsf{Kn}^2)$$

with ρ_g satisfying

$$\begin{cases} -\partial_t\rho_g = C\Delta_x\rho_g - \sigma_a\rho_g \quad \text{in} \quad (0,T)\times\Omega\,, \\ \rho_g(T,x) = 0 \quad \text{on} \quad \{t=T\}\times\Omega\,, \\ \rho_g(t,x) = \eta_\delta(t,x) \quad \text{on} \quad (0,T)\times\partial\Omega\,, \end{cases} \tag{2.22}$$

where $\eta_\delta(t,x)$ is again computed from the boundary layer equation (2.9) with incoming data specified by $\delta(\tau,y)$, as presented in Remark 2.2.

We plug the expansion of f_0 and g into the definition of γ_{Kn}; then in the interior away from the layer,

$$\begin{aligned} (\gamma_{\mathsf{Kn}})_{\text{int}} &:= -\mathsf{Kn}\int_{\mathbb{S}^{n-1}}\int_0^T f_0(t,x,v;\phi)g(t,x,v;\tau,y)\mathrm{d}t\mathrm{d}v\,, \\ &= -\mathsf{Kn}\int_0^T \rho_f(t,x)\rho_g(t,x)\mathrm{d}t + \mathcal{O}(\mathsf{Kn}^3)\,. \end{aligned}$$

Here in the derivation, the $\mathcal{O}(\mathsf{Kn}^2)$ terms

$$\int_0^T \rho_f\int_{\mathbb{S}}^{n-1} v\cdot\nabla_x\rho_g\mathrm{d}v\mathrm{d}t + \int_0^T\rho_g\int_{\mathbb{S}}^{n-1} v\cdot\nabla_x\rho_f\mathrm{d}v\mathrm{d}t$$

disappear since the integrands are odd functions. Simplification is not available inside the boundary layer. Then inserting this γ_{Kn} back into (2.19) we have

$$\begin{aligned} \int_\Omega c(x)\gamma_{\mathsf{Kn}}(x)\mathrm{d}x = &-\mathsf{Kn}\int_{\text{int}} c(x)\int_0^T \rho_f(t,x)\rho_g(t,x)\mathrm{d}t\mathrm{d}x \\ &+\int_{\text{bl}} c(x)\int_0^T \gamma_{\mathsf{Kn}}(x)\mathrm{d}t\mathrm{d}x + \mathcal{O}(\mathsf{Kn}^3)\,. \end{aligned}$$

Since $c(x)$ vanishes inside the layer, we have $c(x)\sim\mathcal{O}\left(\frac{\delta}{\mathsf{Kn}}\right)$. Recalling the definition of κ_a in (2.18) and the fact that $\|\tilde\sigma_a\|_{L^\infty}\sim\mathcal{O}(1)$, the result readily follows. ∎

We note that ρ_f and ρ_g are solutions to the heat equation and have no dependence on Kn. In the time-independent case [14], we show further that the product of ρ_f and ρ_g, the solutions to two elliptic equations, are of low rank, so that we get a better estimate on κ_a ($\mathcal{O}(\frac{\delta}{\mathsf{Kn}^2})$ instead of $\mathcal{O}(\frac{\delta}{\mathsf{Kn}})$). However, this is not true for the time-dependent case: we cannot prove the term $\int_0^T\rho_f\rho_g\mathrm{d}t$ being low rank in $\mathbb{L}^2$, and thus it is hard to obtain $\mathcal{O}(\delta/\mathsf{Kn}^2)$.

Remark 2.6. We would like to point out that the definition of distinguishability coefficient (2.18) is a "continuous" analogue of the condition number in the discrete setting. In fact, if we discretize (2.17) in x and write it in a matrix form, we get

$$\mathsf{A}\cdot\tilde\sigma_a^{dis} = \mathsf{b}\,,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where each row of A is γ_{Kn} evaluated at all discrete points with one particular τ, y, and ϕ selected:

$$\mathsf{A}_{ij} = \gamma_{\mathsf{Kn}}(x_j; \tau_i, y_i, \phi_i) \quad \text{and} \quad \mathsf{b}_i = b(\tau_i, y_i, \phi_i)\,.$$

Perform the singular value decomposition of A

$$\mathsf{A} = \mathsf{U}\cdot\Sigma\cdot\mathsf{V}^T = \sum_{i=1}^{N} \lambda_i \mathsf{u}_i \mathsf{v}_i^T\,, \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N\,,$$

with λ_i being the singular values and u_i and v_i the column vectors; then

$$\tilde{\sigma}_a^{dis} = \mathsf{V}\cdot\Sigma^{-1}\cdot\mathsf{U}^T\mathsf{b}\,.$$

Similarly, a variation of $\tilde{\sigma}_a^{dis}$, denoted as σ_a^{dis}, satisfies

$$\sigma_a^{dis} = \mathsf{V}\cdot\Sigma^{-1}\cdot\mathsf{U}^T(\mathsf{b}+\mathsf{b}^\delta)\,.$$

Then the equivalent definition of κ here is

$$\kappa_{\mathsf{A}} = \max_{\mathsf{b}^\delta:\|\mathsf{b}^\delta\|_\infty<\delta} \frac{\|\sum \frac{1}{\lambda_i}\mathsf{v}_i\mathsf{u}_i^T\mathsf{b}^\delta\|_\infty}{\|\sum \frac{1}{\lambda_i}\mathsf{v}_i\mathsf{u}_i^T\mathsf{b}\|_\infty}\,.$$

Assume b is a fixed vector, and let the denominator be $\mathcal{O}(1)$; then the biggest number is achieved if b^δ is aligned with u_N so that $\kappa_{\mathsf{A}} = \frac{\delta}{\lambda_N}$. Or κ_{A} may implicitly depend on b as well, and if the definition is replaced by

$$\kappa_{\mathsf{A}} = \max_{\mathsf{b}^\delta,\mathsf{b}:\|\mathsf{b}^\delta\|_\infty<\delta\|\mathsf{b}\|_\infty} \frac{\|\sum \frac{1}{\lambda_i}\mathsf{v}_i\mathsf{u}_i^T\mathsf{b}^\delta\|_\infty}{\|\sum \frac{1}{\lambda_i}\mathsf{v}_i\mathsf{u}_i^T\mathsf{b}\|_\infty}\,,$$

then the maximum is achieved by the condition number, $\kappa_{\mathsf{A}} = \frac{\delta\lambda_1}{\lambda_N}$, by aligning b with u_1 and b^δ with u_N.

2.3. Recover scattering coefficient $\boldsymbol{\sigma_s}$. To recover σ_s we follow the same route: first set up the inverse problem through a linearization and show that the stability degrades as $\mathsf{Kn}\to 0$. Without loss of generality, we set $\sigma_a = 1$.

2.3.1. Inverse problem setup. To set up the inverse problem we first recall the forward problem:

$$\begin{cases} \mathsf{Kn}\partial_t f + v\cdot\nabla_x f = \frac{1}{\mathsf{Kn}}\sigma_s\mathcal{L}f - \mathsf{Kn}f \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,, \\ f(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{n-1}\,, \\ f(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,; \end{cases} \tag{2.23}$$

then a similar linearization procedure can be conducted as follows. Assume that $\sigma_s(x)$ can be written as a superposition of a known background $\sigma_{s0}(x)$ and a perturbation $\tilde{\sigma}_s(x)$ from the background, i.e.,

$$\sigma_s(x) = \tilde{\sigma}_s(x) + \sigma_{s0}(x) \quad \text{with} \quad |\tilde{\sigma}_s| \ll |\sigma_s|\,, \quad \text{a.e.}\,;$$

then the background solution f_0 satisfies the equation

$$\begin{cases} \mathsf{Kn}\partial_t f_0 + v\cdot\nabla_x f_0 = \frac{1}{\mathsf{Kn}}\sigma_{s0}\mathcal{L}f_0 - \mathsf{Kn}f_0 \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,, \\ f_0(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{n-1}\,, \\ f_0(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,. \end{cases} \tag{2.24}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The residue $\tilde{f} := f - f_0$ then solves

$$\begin{cases} \mathsf{Kn}\partial_t \tilde{f} + v\cdot\nabla_x \tilde{f} = \frac{1}{\mathsf{Kn}}\sigma_{s0}\mathcal{L}\tilde{f} - \frac{1}{\mathsf{Kn}}\tilde{\sigma}_s\mathcal{L}f_0 - \mathsf{Kn}\tilde{f} \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,, \\ \tilde{f}(0,x,v) = 0 \quad \text{on} \quad \{t=0\}\times\Omega\times\mathbb{S}^{n-1}\,, \\ \tilde{f}(t,x,v) = 0 \quad \text{on} \quad (0,T)\times\Gamma_-\,. \end{cases} \tag{2.25}$$

Write the adjoint problem of (2.24) as

$$\begin{cases} -\mathsf{Kn}\partial_t g - v\cdot\nabla_x g = \frac{1}{\mathsf{Kn}}\sigma_{s0}\mathcal{L}g - \mathsf{Kn}g \quad \text{in} \quad (0,T)\times\Omega\times\mathbb{S}^{n-1}\,, \\ g(T,x,v) = 0 \quad \text{on} \quad \{t=T\}\times\Omega\times\mathbb{S}^{n-1}\,, \\ g(t,x,v) = \delta(\tau,y) \quad \text{on} \quad (0,T)\times\Gamma_+\,, \end{cases} \tag{2.26}$$

then multiply it with $\tilde{f}$ and subtract the product of (2.25) with g, and integrate over $(0,T)\times\Omega\times\mathbb{S}^{n-1}$; we get

(2.27)
$$\int_{\Gamma_+(y)} \tilde{f}(\tau,y,v)n(y)\cdot v\mathrm{d}v = \frac{1}{\mathsf{Kn}}\int_\Omega \tilde{\sigma}_s(x)\int_{\mathbb{S}^{n-1}}\int_0^T g(t,x,v;\tau,y)\mathcal{L}f_0(t,x,v;\phi)\mathrm{d}t\mathrm{d}v\mathrm{d}x\,.$$

Consequently, it prompts a linear equation for $\tilde{\sigma}_a$,

$$\int_\Omega \tilde{\sigma}_s(x)\gamma_{\mathsf{Kn}}(x;\tau,y,\phi)\mathrm{d}x = b(\tau,y,\phi)\,, \tag{2.28}$$

where

$$\begin{aligned} \gamma_{\mathsf{Kn}}(x;\tau,y,\phi) &:= \frac{1}{\mathsf{Kn}}\int_{\mathbb{S}^{n-1}}\int_0^T g(t,x,v;\tau,y)\mathcal{L}f_0(t,x,v;\phi)\mathrm{d}t\mathrm{d}v \\ &= \frac{1}{\mathsf{Kn}}\int_0^T \langle g\rangle_v\langle f\rangle_v - \langle gf\rangle_v\,dt, \end{aligned} \tag{2.29}$$

and

$$b(\tau,y,\phi) := \int_{\Gamma_+(y)} (f(\tau,y,v) - f_0(\tau,y,v))n(y)\cdot v\mathrm{d}v\,.$$

Here $\langle f\rangle_v := \frac{1}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{n-1}} f(t,x,v)\mathrm{d}v$. Again $b(\tau,y,\phi)$ is the data at our disposal—the difference between the measured data and computed data—and we end up with a Fredholm operator of the first kind with kernel $\gamma_{\mathsf{Kn}}(x)$. We study in the next section the ill-conditioning for this family of linear equations in the diffusion regime.

Remark 2.7. In deriving (2.25), we have assumed that $\frac{1}{\mathsf{Kn}}|\tilde{\sigma}_s\tilde{f}| \ll 1$ so that it can be omitted, which restricts the magnitude of the allowable perturbation. We leave the more general perturbation and fully nonlinear inverse problem to the future work.

2.3.2. Ill-conditioning in the diffusion limit. Similar to the previous case, when Kn decreases, the transport equation approaches a diffusion equation and thus recovering the scattering coefficient σ_s is less stable. More precisely, we have the following theorem.

THEOREM 2.8. *For a family of linear equations defined in* (2.28) *and an error tolerance* $\delta > 0$ *on the measurement, define the distinguishability coefficient as*

$$\kappa_s := \sup_{\hat{\sigma}_s\in\Gamma_\delta} \frac{\|\hat{\sigma}_s - \tilde{\sigma}_s\|_{L^\infty(\mathrm{d}x)}}{\|\tilde{\sigma}_s\|_{L^\infty(\mathrm{d}x)}}\,, \tag{2.30}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

$$\Gamma_\delta = \{\hat{\sigma}_s : \sup_{\substack{\forall \|\phi\|_{L^\infty(\Gamma_-)} \le 1, \\ \forall y \in \partial\Omega,\ \tau \in [0,T]}} |\langle \gamma_{\mathsf{Kn}}, \sigma_s \rangle_{L^2(\mathrm{d}x)} - b(\tau, y, \phi)| \le \delta\},$$

and $\tilde{\sigma}_s$ *is the exact solution to* (2.28). *Then we have*

$$\kappa_s \ge \mathcal{O}\left(\frac{\delta}{\mathsf{Kn}}\right) \quad \textit{when} \quad \mathsf{Kn} \ll 1.$$

Proof. The proof again follows a boundary-interior decomposition and asymptotic expansion. First write g and f_0 as

$$f_0 = f_{\mathrm{bl}} + f_{\mathrm{int}}, \quad g = g_{\mathrm{bl}} + g_{\mathrm{int}},$$

where f_{bl} and g_{bl} are the boundary layer part, and f_{int} and g_{int} are the interior part that admits the following expansion:

$$\begin{aligned} f_{\mathrm{int}} &= \rho_f - \mathsf{Kn}\frac{v \cdot \nabla_x \rho_f}{\sigma_{s0}} + \mathsf{Kn}^2 f_2, \\ g_{\mathrm{int}} &= \rho_g + \mathsf{Kn}\frac{v \cdot \nabla_x \rho_g}{\sigma_{s0}} + \mathsf{Kn}^2 g_2. \end{aligned} \tag{2.31}$$

Here ρ_f and ρ_g satisfy the diffusion equations:

$$\partial_t \rho_f + \rho_f - C\nabla_x\left(\frac{1}{\sigma_{s0}}\nabla_x \rho_f\right) = 0, \quad \partial_t \rho_g - \rho_g + C\nabla_x\left(\frac{1}{\sigma_{s0}}\nabla_x \rho_g\right) = 0$$

with suitable initial data and boundary condition.

Now decompose γ_{Kn} also into a layer and interior parts, i.e., $\gamma_{\mathsf{Kn}}(x) = (\gamma_{\mathsf{Kn}}(x))_{\mathrm{bl}} + (\gamma_{\mathsf{Kn}}(x))_{\mathrm{int}}$, where $(\gamma_{\mathsf{Kn}}(x))_{\mathrm{int}}$ is supported on the interior domain and $(\gamma_{\mathsf{Kn}}(x))_{\mathrm{bl}}$ is on the boundary layer. Then for the interior part, according to the definition (2.29), using (2.31), we have

$$\begin{aligned} (\gamma_{\mathsf{Kn}})_{\mathrm{int}} &= \frac{1}{\mathsf{Kn}}\int_0^T \left(\langle f_{\mathrm{int}}\rangle_v \langle g_{\mathrm{int}}\rangle_v - \langle f_{\mathrm{int}} g_{\mathrm{int}}\rangle_v\right) \mathrm{d}t \\ &= \frac{1}{\mathsf{Kn}}\int_0^T \left(\rho_f + \mathsf{Kn}^2 \langle f_2\rangle_v\right)\left(\rho_g + \mathsf{Kn}^2 \langle g\rangle_v\right) \\ &\qquad - \left\langle \left(\rho_f - \frac{\mathsf{Kn}}{\sigma_{s0}} v\cdot\nabla_x\rho_f + \mathsf{Kn}^2 f_2\right)\left(\rho_g + \frac{\mathsf{Kn}}{\sigma_{s0}} v\cdot\nabla_x\rho_g + \mathsf{Kn}^2 g_2\right)\right\rangle_v \mathrm{d}t \\ &= \frac{\mathsf{Kn}}{\sigma_{s0}^2}\int_0^T \langle (v\cdot\nabla_x\rho_f)(v\cdot\nabla_x\rho_g)\rangle_v \,\mathrm{d}t + \mathcal{O}(\mathsf{Kn}^2) \\ &= \frac{C\mathsf{Kn}}{\sigma_{s0}^2}\int_0^T \nabla_x\rho_f\cdot\nabla_x\rho_g \mathrm{d}t + \mathcal{O}(\mathsf{Kn}^2), \end{aligned} \tag{2.32}$$

where C again depends on the dimension of the velocity space. Here in the third equality, the terms $\langle \frac{\mathsf{Kn}}{\sigma_s}(v\cdot\nabla_x\rho_f)\rho_g\rangle_v$ and $\langle \frac{\mathsf{Kn}}{\sigma_s}(v\cdot\nabla_x\rho_g)\rho_f\rangle_v$ vanish as the integrands are odd.

Now choose $c(x)$ such that it vanishes inside the boundary layer and has

$$|\langle \gamma_{\mathsf{Kn}}, c\rangle_{L^2(\mathrm{d}x)}| \le \delta,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

then $\hat{\sigma}_s$ constructed as $\hat{\sigma}_s(x) = c(x) + \tilde{\sigma}_s(x) \in \Gamma_\delta$. Computing

$$\langle \gamma_{\mathsf{Kn}}, c\rangle_{L^2(\mathrm{d}x)} = \langle (\gamma_{\mathsf{Kn}})_{\mathrm{int}}, c_{\mathrm{int}}\rangle_{L^2(\mathrm{d}x)} = \left\langle -\frac{C\mathsf{Kn}}{\sigma_{s0}^2}\int_0^T \nabla_x \rho_f \cdot \nabla_x \rho_g \mathrm{d}t,\, c_{\mathrm{int}} \right\rangle_{L^2(\mathrm{d}x)} + \mathcal{O}(\mathsf{Kn}^2)\,,$$

we see that

$$c(x) \sim \mathcal{O}\left(\frac{\delta}{\mathsf{Kn}}\right).$$

Again, from the definition of κ_s and the fact that $\|\tilde{\sigma}_s\|_{L^\infty} \sim \mathcal{O}(1)$, we get $\kappa_s \geq \mathcal{O}\left(\frac{\delta}{\mathsf{Kn}}\right)$. ∎

3. Highly forward peaked regime. In this section, we consider the anisotropic scattering and study the well-/ill-posedness of the inverse RTE in the highly forward peaked regime, in which the time-dependent RTE is asymptotically equivalent to the Fokker–Planck equation. For simplicity, we study the critical case with zero absorption and x-independent scattering. The RTE reads

$$\begin{cases} \partial_t f(t,x,v) + v\cdot\nabla_x f(t,x,v) = \mathcal{L}f(t,x,v)\,, \\ f(0,x,v) = f^I(x,v) \quad \text{on} \quad \Omega\times\mathbb{S}^{d-1}\,, \\ f(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,, \end{cases} \tag{3.1}$$

where the collision operator takes the form

$$\mathcal{L}f(t,x,v) = \frac{1}{\epsilon^2}\int_{\mathbb{S}^{d-1}} \sigma\left(\frac{1-v\cdot v'}{\epsilon}\right)(f(t,x,v') - f(t,x,v))\mathrm{d}v'\,. \tag{3.2}$$

Without loss of generality, we assume that σ integrates to one, i.e.,

$$\frac{1}{\epsilon}\int_{\mathbb{S}^{d-1}} \sigma\left(\frac{1-v\cdot v'}{\epsilon}\right)\mathrm{d}v' = 1\,. \tag{3.3}$$

Considering v' is the incident direction and v is the scattering direction, then the small parameter ϵ reinforces "small-angle" scattering—the kernel is peaked in the forward direction of flight; it also plays a role of mean free path, which accounts for the strong scattering effect. Here $v \in \mathbb{S}^{d-1}$ is a unit vector denoting the direction of flight. Hereafter, we will focus on dimension $d = 3$.

The existence of such a regime was long exposed to the area [28, 24] but has received little attention in the inverse problem setting. It is not quite known how stabilities change according to ϵ despite some conjectures [5]. We address this issue in this section. We will first formally derive the Fokker–Planck limit in section 3.1 and set up the inverse problem in section 3.2. Stability with respect to ϵ will be discussed in section 3.3.

3.1. Fokker–Planck limit. The equation, in the zero limit of ϵ, loses the large-angle scattering and effectively is equivalent to the Fokker–Planck equation. The original derivation was seen in [28, 24]. Denote $\mu = v\cdot v'$ the cosine of the scattering angle; then the scattering cross section has the Legendre polynomial expansion,

$$\frac{1}{\epsilon}\sigma\left(\frac{1-\mu}{\epsilon}\right) = \sum_{n=0}^{\infty}\frac{2n+1}{4\pi}\sigma_n P_n(\mu)\,,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where the projection coefficients on the nth Legendre polynomial $P_n(\mu)$ are

$$\sigma_n = \frac{2\pi}{\epsilon}\int_{-1}^{1}\sigma\left(\frac{1-\mu}{\epsilon}\right)P_n(\mu)\mathrm{d}\mu\,. \tag{3.4}$$

It is immediate that $\sigma_0 = 1$ from (3.3).

To proceed, we write v using spherical coordinates, $v = (\sqrt{1-v_3^2}\cos\psi, \sqrt{1-v_3^2}\sin\psi, v_3)$, and introduce the spherical harmonic functions

$$Y_{n,m}(v) = \left[\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}\right]^{1/2}\times(-1)^{(m+|m|)/2}P_{n,|m|}(v_3)e^{in\psi}\,, n\geq 0,\ \ -n\leq m\leq n\,,$$

where $P_{n,|m|}(v_3)$ are associated Legendre functions

$$P_{n,|m|}(v_3) = (1-v_3^2)^{m/2}\left(\frac{\mathrm{d}}{\mathrm{d}v_3}\right)^m P_n(v_3),\quad 0\leq m\leq n\,.$$

The spherical harmonic functions form a complete set of orthonormal basis and thus any suitably smooth function $f(v)$ defined on the unit sphere can be expanded as

$$f(v) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} f_{n,m}Y_{n,m}(v),\qquad f_{n,m} := \int_{\mathbb{S}^2} f(v)Y_{n,m}(v)\mathrm{d}v\,. \tag{3.5}$$

Note also that $P_n(\mu)$ satisfy the addition formula:

$$P_n(v\cdot v') = \frac{4\pi}{2n+1}\sum_{m=-n}^{n} Y_{n,m}(v)Y_{n,m}(v')\,.$$

Therefore, the collision (3.2) admits the following expansion:

$$\begin{aligned}\mathcal{L}f &= \frac{1}{\epsilon}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\sigma_n\int_{\mathbb{S}^2}Y_{n,m}(v)Y_{n,m}(v')\left[f(v')-f(v)\right]\mathrm{d}v'\\ &= \frac{1}{\epsilon}\left[\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\sigma_n f_{n,m}Y_{n,m}(v)-\sigma_0 f(v)\right]\\ &= \frac{1}{\epsilon}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}(\sigma_n-\sigma_0)f_{n,m}Y_{n,m}(v)\,.\end{aligned} \tag{3.6}$$

The second equality holds because $\int_{\mathbb{S}^2}Y_{n,m}\mathrm{d}v = 0$ for all $n\geq 1$ and $\int_{\mathbb{S}^2}Y_{n,m}\mathrm{d}v = 1$ only if $n=m=0$. Letting $\alpha = \frac{1-\mu}{\epsilon}$, we rewrite σ_n in (3.4) as

$$\begin{aligned}\sigma_n &= 2\pi\int_0^{2/\epsilon}\sigma(\alpha)P_n(1-\epsilon\alpha)\mathrm{d}\alpha\\ &= 2\pi\int_0^{2/\epsilon}\sigma(\alpha)\left[P_n(1)-P_n'(1)\epsilon\alpha+\frac{P_n''(1)}{2}(\epsilon\alpha)^2+\cdots\right]\mathrm{d}\alpha\,.\end{aligned} \tag{3.7}$$

If we define

$$\begin{aligned}\xi_n &:= 2\pi\int_{-1}^{1}\frac{1}{\epsilon^2}\sigma\left(\frac{1-\mu}{\epsilon}\right)(1-\mu)^n\mathrm{d}\mu\\ &= \epsilon^{n-1}\left(2\pi\int_0^{2/\epsilon}t^n\sigma(t)\mathrm{d}t\right)\\ &= \mathcal{O}(\epsilon^{n-1})\,,\end{aligned} \tag{3.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

then σ_n can be rewritten as

$$\sigma_n = \epsilon \left[P_n(1)\xi_0 + P_n'(1)\xi_1 + \frac{1}{2}P_n''(1)\xi_2 + \frac{1}{3!}P_n'''(1)\xi_3 + \cdots \right]. \tag{3.9}$$

Note that ξ_0 is fixed and has no dependence on σ due to (3.3):

$$\xi_0 = 2\pi \int_{-1}^{1} \frac{1}{\epsilon^2}\sigma\left(\frac{1-\mu}{\epsilon}\right) \mathrm{d}\mu = \frac{2\pi}{\epsilon}. \tag{3.10}$$

Since

$$P_n(1) = 1, \quad P_n'(1) = \frac{n(n+1)}{2}, \quad P_0(\mu) = 1,$$

we have from (3.7) that

$$\sigma_n - \sigma_0 = -\epsilon \frac{n(n+1)}{2}\xi_1 + \mathcal{O}(\epsilon^2),$$

and therefore plugging it into (3.6) we get

$$\mathcal{L}f = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} -\frac{n(n+1)}{2}\xi_1 f_{n,m} Y_{n,m} + \mathcal{O}(\epsilon). \tag{3.11}$$

Recall that for the well-known Fokker–Planck operator in spherical coordinates,

$$\mathcal{L}_{\mathrm{FP}} f(v) = \left[\frac{\partial}{\partial v_3}(1-v_3^2)\frac{\partial}{\partial v_3} + \frac{1}{1-v_3^2}\frac{\partial^2}{\partial \psi^2} \right] f(v),$$

we have

$$\mathcal{L}_{\mathrm{FP}} Y_{n,m}(v) = -n(n+1) Y_{n,m}(v). \tag{3.12}$$

Comparing (3.11) and (3.12), we get the Fokker–Planck approximation:

$$\mathcal{L}f(v) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} -\left(\frac{n(n+1)}{2}\xi_1 + \mathcal{O}(\epsilon) \right) f_{n,m} Y_{n,m} \quad = \frac{\xi_1}{2}\mathcal{L}_{\mathrm{FP}} f(v) + \mathcal{O}(\epsilon)$$

with

$$\xi_1 = 2\pi \int_0^{2/\epsilon} t\sigma(t)\mathrm{d}t \sim 2\pi \int_0^{\infty} t\sigma(t)\mathrm{d}t. \tag{3.13}$$

In other words, when ϵ is small, the linear scattering operator $\mathcal{L}$ converges to the Fokker–Planck operator with a scalar multiplication and the linear transport equation converges to the Fokker–Planck equation

$$\partial_t f + v \cdot \nabla_x f = \frac{\xi_1}{2}\mathcal{L}_{\mathrm{FP}} f,$$

where $\mathcal{L}_{\mathrm{FP}}$ and ξ_1 are defined in (3.12) and (3.13), respectively.

Remark 3.1. We note that the unknown in the collision term defined in (3.2) is $\sigma(\mu)$. As a function of μ, it could be fully recovered only if all the coefficients σ_n in (3.4) are known. According to (3.9), this requires knowledge about ξ_n for all n. However, in the zero limit, the collision term converges to the Laplace operator, and there is only one scalar that is unknown: ξ_1. As a result, the limiting Fokker–Planck equation is much easier to invert heuristically. This will be reflected in section 3.3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3.2. Inverse problem setup. In the inverse problem setting, we are given inflow data and measure the outflow, with which we infer the scattering coefficient $\sigma(v, v')$. Here the albedo operator is given by

$$\mathcal{A}(\sigma): \qquad \phi(t,x,v)|_{(0,T)\times\Gamma_-} \to \int_{\Gamma_+(y)} f(t,y,v)n(y)\cdot v\mathrm{d}v\,.$$

We first linearize the albedo operator. Like always, we assume that a priori knowledge provides a background state σ_0 such that the residue $\tilde{\sigma} := \sigma - \sigma_0$ satisfies

$$|\tilde{\sigma}| \ll |\sigma|\,, \quad \text{a.e.;}$$

then with background state σ_0, one gets the solution f_0 that solves the following initial boundary value problem:

$$\begin{cases} \partial_t f_0(t,x,v) + v\cdot\nabla_x f_0(t,x,v) = \mathcal{L}_0 f_0(t,x,v)\,, \\ f_0(0,x,v) = 0 \quad \text{on} \quad \Omega\times\mathbb{S}^2\,, \\ f_0(t,x,v) = \phi(t,x,v) \quad \text{on} \quad (0,T)\times\Gamma_-\,, \end{cases} \tag{3.14}$$

where

$$\mathcal{L}_0 f_0(t,x,v) = \frac{1}{\epsilon^2}\int_{4\pi}\sigma_0\left(\frac{1-\mu}{\epsilon}\right)(f_0(t,x,v') - f_0(t,x,v))\mathrm{d}v'\,.$$

The residue

$$\tilde{f}(t,x,v) := f(t,x,v) - f_0(t,x,v)$$

then satisfies

$$\partial_t\tilde{f}(t,x,v) + v\cdot\nabla_x\tilde{f}(t,x,v) = \mathcal{L}_0\tilde{f}(t,x,v) + \tilde{\mathcal{L}}f_0(t,x,v) \tag{3.15}$$

with zero initial data and boundary data. Here

$$\tilde{\mathcal{L}}f_0(t,x,v) = \frac{1}{\epsilon^2}\int_{4\pi}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right)(f_0(t,x,v') - f_0(t,x,v))\mathrm{d}v'\,. \tag{3.16}$$

We also define an adjoint problem to (3.14):

$$\begin{cases} -\partial_t g(t,x,v) - v\cdot\nabla_x g(t,x,v) = \mathcal{L}_0 g(t,x,v)\,, \\ g(T,x,v) = 0 \quad \text{on} \quad \Omega\times\mathbb{S}^2\,, \\ g(t,x,v) = \delta(\tau,y) \quad \text{on} \quad (0,T)\times\Gamma_+\,. \end{cases} \tag{3.17}$$

Multiply (3.15) and (3.17) by g and $\tilde{f}$, respectively, and subtract them; we get, after integrating in x, v, and t,

$$\int_{\Gamma_+(y)}\tilde{f}(\tau,y,v)n(y)\cdot v\mathrm{d}v = \int_{\Omega\times\mathbb{S}^2}\int_0^T g(t,x,v)\tilde{\mathcal{L}}f_0(t,x,v)\mathrm{d}t\mathrm{d}v\mathrm{d}x\,, \tag{3.18}$$

where the LHS is the difference between the measurement of $f(t,x,v)$ and the computed $f_0(t,x,v)$ at time τ and position $y\in\partial\Omega$, and we denote it by $b(\tau,y,\phi)$,

$$b(\tau,y,\phi) = \int_{\Gamma_+(y)}\tilde{f}(\tau,y,v)n(y)\cdot v\mathrm{d}v\,.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The RHS of (3.18) gives a linear function for $\tilde{\sigma}$. In particular, using (3.16) we have

$$b(\tau, y, \phi) = \frac{1}{\epsilon^2} \int_{\mathbb{S}^2 \times \mathbb{S}^2} \tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) \gamma_\epsilon(v, v') \mathrm{d}v' \mathrm{d}v \,, \tag{3.19}$$

where

$$\gamma_\epsilon(v, v'; \tau, y, \phi) := \int_0^T \int_\Omega g(t, x, v)[f_0(t, x, v') - f_0(t, x, v)] \mathrm{d}x \mathrm{d}t \,. \tag{3.20}$$

By varying $\{\tau, y\}$ and ϕ, one obtains different g and f_0, and thus γ_ϵ, making (3.19) a Fredholm operator of the first kind with parameters $\{\tau, y, \phi\}$.

As mentioned in Remark 2.3, to have a unique recovery of $\tilde{\sigma}$ in L_p space, one needs γ_ϵ expanding the adjoint space L_q (with $\frac{1}{p} + \frac{1}{q} = 1$). The injectivity is beyond the scope of the current paper, and we only discuss the stability in the following section.

3.3. Stability in the highly forward peaked regime. This subsection is devoted to the stability in the recovery of σ in the forward peaked regime. There are two aspects of the problem:

1. To fully recover $\sigma(1 - v \cdot v')$, as mentioned in Remark 3.1, one needs all its moments σ_n, which in turn requires the information of ξ_n for all n. However, since ξ_n diminishes at the order of ϵ^{n-1}, obtaining ξ_n is very sensitive to the pollution in the data. Indeed, suppose the data has pollution of order δ; then there are at most $n_0 = \log_\epsilon \delta + 1$ terms that can be recovered. Now keeping δ fixed and sending ϵ to 0, the number n_0 decreases to 1, meaning that all the higher order information gets lost. This is indeed consistent with the view of the singular decomposition [17, 5]. In that viewpoint, the reconstruction of σ relies on the separation of the ballistic component (pure transport) and the scattered components (mainly the single-scattering). In the forward peaked regime, however, the single-scattering concentrates on the original velocity and does not distinguish from the ballistic transport much, making the separation hard, and thus disables the reconstruction. This will be demonstrated in Theorem 3.3.
2. Nevertheless, in the Fokker–Planck regime, it's not the full information $\sigma(1 - v \cdot v')$ that matters but the rescaled one $\frac{1}{\epsilon}\sigma(\frac{1-v\cdot v'}{\epsilon})$. As written, when ϵ is small, the rescaled σ will concentrate around $v \cdot v' = 1$ and only a little information is needed to recover its shape. Indeed, according to (3.7), (3.8), and (3.9), ξ_n quickly decays to zero, and all of σ_n are dominated by the first few ξ_n for certain accuracy. For example, if ϵ accuracy is needed for σ_n, one only needs to recover one parameter ξ_1. This significantly reduces the amount of measurements needed. In this sense, we find that the inverse problem with highly forward peaked scattering is actually more practice friendly. This is demonstrated in Theorem 3.2.

To recover $\frac{1}{\epsilon}\tilde{\sigma}_s(\frac{1-v\cdot v'}{\epsilon})$, one simply needs to find all its Legendre coefficients $\tilde{\sigma}_n$ in the expansion

$$\frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1 - v \cdot v'}{\epsilon}\right) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \tilde{\sigma}_n P_n(\mu) \,, \qquad \mu = v \cdot v' \,. \tag{3.21}$$

Using the same expression as in (3.9), one has

$$\tilde{\sigma}_n = \epsilon \left[P_n(1)\tilde{\xi}_0 + P_n'(1)\tilde{\xi}_1 + \frac{1}{2}P_n''(1)\tilde{\xi}_2 + \frac{1}{3!}P_n'''(1)\tilde{\xi}_3 + \cdots \right] \tag{3.22}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with

$$\tilde{\xi}_n = \epsilon^{n-1} 2\pi \int_0^{\epsilon/2} t^n \tilde{\sigma}_s(t) \mathrm{d}t = \mathcal{O}(\epsilon^{n-1}) \,. \tag{3.23}$$

Introducing the above relations into (3.19), we get

$$\begin{aligned} b(\tau, y, \phi) &= \frac{1}{\epsilon} \sum_{n=0}^{\infty} \tilde{\sigma}_{sn} \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} \xi_j P_n^{(j)}(1) \frac{1}{j!} \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' \\ &= \sum_{j=0}^{\infty} \xi_j \left(\frac{1}{j!} \sum_{n=0}^{\infty} P_n^{(j)}(1) \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' \right) . \end{aligned}$$

Consequently, we obtain the following linear system for $\xi = (\xi_1, \xi_2, \cdots)$:

$$\mathsf{A}\xi = \mathbf{b} \,, \tag{3.24}$$

where $\mathbf{b}$ is a column vector whose size is equal to the number of experiments, and in matrix $\mathsf{A} = [a_{ij}]$, the component a_{ij} is determined by

$$a_{ij} = \frac{1}{j!} \sum_{n=0}^{\infty} P_n^{(j)}(1) \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' \,,$$

where the subscript i represent experiment i with choice τ_i, y_i, ϕ_i. Recalling the expression of γ_ϵ in (3.20), one has

$$\begin{aligned} \int_{\mathbb{S}^2} P_n(\mu) \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' &= \frac{4\pi}{2n+1} \sum_{m=-n}^{n} \int Y_{n,m}(v) Y_{n,m}(v') \gamma_\epsilon(v, v') \mathrm{d}v \mathrm{d}v' \\ &= \frac{4\pi}{2n+1} \sum_{m=-n}^{n} \int Y_{n,m}(v) Y_{n,m}(v') \\ &\quad \times \left[\int g(v) f_0(v') - g(v) f_0(v) \mathrm{d}x \mathrm{d}t \right] \mathrm{d}v \mathrm{d}v' \\ &= \frac{4\pi}{2n+1} \sum_{m=-n}^{n} \left[\bar{g}_{n,m} (\overline{f_0})_{n,m} - (\overline{g f_0})_{n,m} \delta_{n,m} \right] , \end{aligned}$$

where the over-line denotes integration in both x and t. Therefore,

$$a_{ij} = \sum_{n=0}^{\infty} \frac{P_n^{(j)}}{j!} \frac{4\pi}{2n+1} \sum_{m=-n}^{n} \left[\bar{g}_{n,m} (\overline{f_0})_{n,m} - (\overline{g f_0})_{n,m} \delta_{n,m} \right] . \tag{3.25}$$

THEOREM 3.2. *The recovery of* $\frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right)$ *does not deteriorate as* $\epsilon \to 0$. *More precisely, if we define the distinguishability coefficient as*

$$\kappa_\epsilon = \sup_{\hat{\sigma} \in \Gamma_\delta} \frac{\| \frac{1}{\epsilon}\hat{\sigma}\left(\frac{1-\mu}{\epsilon}\right) - \frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) \|_\infty}{\left\| \frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) \right\|_\infty} \,, \tag{3.26}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

$$\Gamma_\delta = \left\{ \frac{1}{\epsilon}\hat{\sigma}\left(\frac{1-\mu}{\epsilon}\right) : \sup_{\substack{\forall \|\phi\|_{L^\infty(\Gamma_-)} \le 1, \\ \forall y \in \partial\Omega,\ \tau \in [0,T]}} \left| \frac{1}{\epsilon} \int \times \left[\frac{1}{\epsilon}\hat{\sigma}\left(\frac{1-\mu}{\epsilon}\right) - \frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) \right] \gamma_\epsilon(v, v'; \tau, y, \phi) \mathrm{d}v\mathrm{d}v' \right| \le \delta \right\},$$

then for $\epsilon \ll \delta$

$$\kappa_\epsilon \sim \mathcal{O}(\delta\epsilon)\,.$$

Proof. From (3.21)–(3.23), we see that

$$\frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) = \epsilon \sum_{k=0}^{\infty} \left[\sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\mu) \frac{P_n^{(k)}(1)}{k!} \right] \tilde{\xi}_k := \epsilon \sum_{k=0}^{\infty} h_k(\mu)\tilde{\xi}_k\,,$$

where $\tilde{\xi}_k$ is defined the same as in (3.23). Similarly, $\frac{1}{\epsilon}\sigma\left(\frac{1-\mu}{\epsilon}\right)$ has the expansion

$$\frac{1}{\epsilon}\sigma\left(\frac{1-\mu}{\epsilon}\right) = \epsilon \sum_{k=0}^{\infty} h_k(\mu)\xi_k\,.$$

Note first that $\tilde{\xi}_0 = \xi_0$; then for $\frac{1}{\epsilon}\sigma(\frac{1-\mu}{\epsilon}) \in \Gamma_\delta$, we have

$$\begin{aligned} &\left| \frac{1}{\epsilon} \int \left[\frac{1}{\epsilon}\sigma\left(\frac{1-\mu}{\epsilon}\right) - \frac{1}{\epsilon}\tilde{\sigma}\left(\frac{1-\mu}{\epsilon}\right) \right] \gamma_\epsilon(v, v'; \tau, y, \phi) \mathrm{d}v\mathrm{d}v' \right| \\ &= \left| \sum_{k=1}^{\infty} \int h_k(\mu)(\xi_k - \tilde{\xi}_k)\gamma_\epsilon \mathrm{d}v\mathrm{d}v' \right| \\ &= \left| \int h_1(\mu)(\xi_1 - \tilde{\xi}_1)\gamma_\epsilon \mathrm{d}v\mathrm{d}v' + \mathcal{O}(\epsilon) \right| \le \delta\,, \end{aligned}$$

which implies that, for $\epsilon \ll \delta$, $|\xi_1 - \tilde{\xi}_1| \le \mathcal{O}(\delta)$ since $h_1(\mu)$ and γ_ϵ are $\mathcal{O}(1)$. Plugging this result into the definition (3.26), one immediately sees that

$$\kappa_\epsilon = \sup_{\sigma \in \Gamma_\delta} \frac{(\xi_1 - \tilde{\xi}_1)h_1(\mu) + \mathcal{O}(\epsilon^2)}{\xi_0 h_0(\mu) + \xi_1 h_1(\mu) + \mathcal{O}(\epsilon^2)} \sim \mathcal{O}(\delta\epsilon)\,,$$

where ϵ on the right comes from the fact that $\xi_0 \sim \mathcal{O}(\epsilon^{-1})$. ∎

Contrary to the above result, if we want to fully recover σ, then the presence of the small scale ϵ will make it impossible. Specifically, we have the following theorem.

THEOREM 3.3. *Suppose* $\sigma \in \mathbb{H}^k(\mathrm{d}\mu)$ *(*$\mathbb{H}^k(\mathrm{d}\mu)$ *is the Sobolev space of functions whose derivatives of order up to* k*-order are in* $\mathbb{L}^2(\mathrm{d}\mu)$*); the recovering* σ *becomes impossible in the limit of* $\epsilon \to 0$ *in the sense that if the* δ *error is allowed in* $\mathbf{b}$ *(e.g., measurement error; see* (3.24)*), then the error in* σ *will be* $\left(\frac{\ln \epsilon}{\ln \delta}\right)^k$.

Proof. Since $\sigma(\mu) \in \mathbb{L}^2(\mathrm{d}\mu)$, we expand it using Hermite functions:

$$\sigma = \sum_n \frac{1}{n!}\hat{\sigma}_n H_n(\mu) \tag{3.27}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with $\hat{\sigma}_n = \langle \sigma\,, H_n \rangle = \int \sigma H_n(\mu)\mathrm{d}\mu$. Here H_n are weighted Hermite functions written as

$$H_n(\mu) = p_n(\mu)e^{-\mu^2/2} \quad \text{with} \quad \int p_m(\mu)p_n(\mu)e^{-\mu^2}\mathrm{d}\mu = \delta_{mn}n!\,.$$

Note that other $\mathbb{L}^2$ basis functions can be used. The Hermite polynomial is only one possible choice.

Meanwhile we recall the definition of ξ:

$$\xi_m = 2\pi\epsilon^{m-1}\int_0^{2/\epsilon}\sigma(\mu)\mu^m\mathrm{d}\mu\,.$$

To prove the theorem, we allow Δ error in σ, and see how much it affects ξ, and then $\mathbf{b}$ in the end. Here $\xi = [\xi_0, \xi_1, \cdots]'$ and $\sigma = [\sigma_0, \sigma_1, \cdots]'$. We first note that $\sigma \in \mathbb{H}^k$, and thus its Hermite polynomial coefficients decay algebraically fast. The standard approximation theory from spectral accuracy indicates

$$\sigma_n = \mathcal{O}(1/n^k)\,.$$

Suppose we tolerate error up to Δ; then one needs to recover σ_n up to at least $n_0 = \Delta^{-1/k}$, and the allowed perturbation in σ_n is

$$\Delta_{\sigma_n} \le \Delta \quad \text{for} \quad n = 0\,, \ldots, n_0\,.$$

We now look for an explicit relation between ξ and σ_n. Considering the explicit relation between monomials and the Hermite polynomials

$$v^m = m!\sum_{k=0}^{m/2}\frac{1}{2^k k!(m-2k)!}p_{m-2k}(v)\,,$$

plugging it back into the equation for ξ_m, we have

$$\begin{aligned}\xi_m &= \epsilon^{m-1}\sum_{k=0}^{m/2}\frac{2\pi m!}{2^k k!(m-2k)!}\int_0^{2/\epsilon}\sigma(v)p_{m-2k}\mathrm{d}v\\ &= \epsilon^{m-1}\sum_{k=0}^{m/2}\sum_{n=0}^{\infty}\frac{2\pi m!}{2^k k!(m-2k)!n!}D_{n,m,k}\sigma_n\,,\end{aligned}$$

where we have used expansion in (3.27) and defined $D_{n,m,k} = \int p_{m-2k}p_n e^{-v^2/2}\mathrm{d}v$. In a matrix form one has

$$\xi = \mathsf{C}\cdot\sigma\,,$$

where C is defined by

$$\mathsf{C}_{mn} = \epsilon^{m-1}\frac{2\pi m!}{n!}\sum_{k=0}^{m/2}\frac{D_{n,m,k}}{2^k k!(m-2k)!}\,.$$

Since we need to recover σ_n up to $n = n_0 = \Delta^{-1/k}$, and the recovered coefficients need to be within error tolerance δ, the tolerance for ξ_{n_0} then is

$$\Delta_{\xi_{n_0}} \le \Delta\sum_{k=0}^{n_0}\mathsf{C}_{n_0,k} \sim \epsilon^{n_0-1}\Delta \quad \text{with} \quad n_0 = \Delta^{-1/k}\,.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Noting the relationship between ξ and $\mathbf{b}$ in (3.24), we see that the error allowance on $\mathbf{b}$ is $\epsilon^{n_0-1}\Delta$. With shrinking ϵ, this restriction becomes more and more severe, making the inverse problem less practical. More specifically if we have δ error in $\mathbf{b}$, then setting

$$\epsilon^{\Delta^{-1/k}}\Delta \sim \delta$$

gives $\Delta > \left(\frac{\ln\epsilon}{\ln\delta}\right)^k \to \infty$ as $\epsilon \to 0$, meaning that the accuracy in recovering σ_n is lost and so it is with σ. ∎

Acknowledgment. The second and third authors are grateful for Prof. Kui Ren's inspiring discussions.

REFERENCES

[1] G. Abdoulaev, K. Ren, and A. Hielscher, *Optical tomography as a PDE-constrained optimization problem*, Inverse Problems, 21 (2005), pp. 1507–1530.

[2] S. Arridge, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), pp. R41–93.

[3] S. Arridge and W. Lionheart, *Nonuniqueness in diffusion-based optical tomography*, Optim. Lett., 23 (1998), pp. 882–884.

[4] S. R. Arridge and J. C. Schotland, *Optical tomography: Forward and inverse problems*, Inverse Problems, 25 (2009), 123010.

[5] G. Bal, *Inverse transport theory and applications*, Inverse Problems, 25 (2009), 053001.

[6] G. Bal, A. Jollivet, I. Langmore, and F. Monard, *Angular average of time-harmonic transport solutions*, Comm. Partial Differential Equations, 36 (2011), pp. 1044–1070.

[7] G. Bal and A. Jollivet, *Time-dependent angularly averaged inverse transport*, Inverse Problems, 25 (2009), 075010.

[8] G. Bal and A. Jollivet, *Stability for time-dependent inverse transport*, SIAM J. Math. Anal., 42 (2010), pp. 679–700.

[9] G. Bal, I. Langmore, and F. Monard, *Inverse transport with isotropic sources and angularly averaged measurement*, Inverse Probl. Imaging, 2 (2008), pp. 23–42.

[10] G. Bal and F. Monard, *Inverse transport with isotropic time-harmonic sources*, SIAM J. Math. Anal., 44 (2012), pp. 134–161.

[11] C. Bardos, S. Santos, and R. Sentis, *Diffusion approximation and computation of the critical size*, Trans. Amer. Math. Soc., 284 (1984), pp. 617–649.

[12] A. Bensoussan, J. L. Lions, and G. C. Papanicolaou, *Boundary layers and homogenization of transport processes*, Publ. Res. Inst. Math. Sci., 15 (1979), pp. 53–157, https://doi.org/10.2977/prims/1195188427.

[13] K. M. Case and P. Zweifel, *Linear Transport Theory*, Addison-Wesley, Reading, MA, 1967.

[14] K. Chen, Q. Li, and L. Wang, *Stability of stationary inverse transport equation in diffusion scaling*, Inverse Problems, 34 (2018), 025004.

[15] Y. Cheng, I. M. Gamba, and K. Ren, *Recovering doping profiles in semiconductor devices with the Boltzmann-Poisson model*, J. Comput. Phys., 230 (2011), pp. 3391–3412.

[16] M. Choulli and P. Stefanov, *Inverse scattering and inverse boundary value problems for the linear Boltzmann equation*, Comm. Partial Differential Equations, 21 (1996), pp. 763–785.

[17] M. Choulli and P. Stefanov, *An inverse boundary value problem for the stationary transport equation*, Osaka J. Math., 36 (1998), pp. 87–104.

[18] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Springer, Berlin, 1993.

[19] A. Greenleaf, M. Lassas, and G. Uhlmann, *On nonuniqueness for Calderón inverse problem*, Math. Res. Lett., 10 (2003), pp. 685–693.

[20] G.Uhlmann, *Electrical impedance tomography and Calderón's problem*, Inverse Problems, 25 (2009), 123011.

[21] I. Langmore, *The stationary transport problem with angularly averaged measurements*, Inverse Problems, 24 (2008), 015024.

[22] E. Larsen, *Solution of three-dimensional inverse transport problems*, Transp. Theory Stat. Phys., 17 (1988), pp. 147–167.

[23] E. Larsen and J. Keller, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys., 15 (1974), pp. 75–81.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[24] C. L. Leakeas and E. W. Larsen, *Generalized Fokker-Planck approximations of particle transport with highly forward-peaked scattering*, Nuclear Sci. Engrg., 137 (2001), pp. 236–250.

[25] M. Machida, G. Y. Panasyuk, Z.-M. Wang, V. A. Markel, and J. C. Schotland, *Radiative transport and optical tomography with large datasets*, J. Opt. Soc. Am. A, 33 (2016), pp. 551–558.

[26] M. Machida and J. C. Schotland, *Inverse Born series for the radiative transport equation*, Inverse Problems, 31 (2015), 095009.

[27] G. Papanicolaou, *Asymptotic analysis of transport process*, Bull. Amer. Math. Soc., 81 (1975), pp. 330–392.

[28] G. Pomraning, *The Fokker-Planck operator as an asymptotic limit*, Math. Models Methods Appl. Sci., 02 (1992), pp. 21–36.

[29] K. Ren, *Recent developments in numerical techniques for transport-based medical imaging methods*, Comm. Comput. Phys, 8 (2010), pp. 1–50.

[30] K. Ren, G. Bal, and A. H. Hielscher, *Transport- and diffusion-based optical tomography in small domains: A comparative study*, Appl. Opt., 46 (2007), pp. 6669–6679.

[31] V. Romanov, *Stability estimates in problems of recovering the attenuation coefficient and the scattering indicatrix for the transport equation*, J. Inverse Ill-Posed Probl., 4 (1996), pp. 297–305.

[32] T. Saratoon, T. Tarvainen, B. T. Cox, and S. R. Arridge, *A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation*, Inverse Problems, 29 (2013), 075006, http://stacks.iop.org/0266-5611/29/i=7/a=075006.

[33] P. Stefanov and A. Tamasan, *Uniqueness and non-uniqueness in inverse radiative transfer*, Proc. Amer. Math. Soc., 137 (2009), pp. 2335–2344.

[34] J. Tang, W. Han, and B. Han, *A theoretical study for RTE-based parameter identification problems*, Inverse Problems, 29 (2013), 095002, http://stacks.iop.org/0266-5611/29/i=9/a=095002.

[35] J. Wang, *Stability estimates of an inverse problem for the stationary transport equation*, Ann. Inst. Henri Poincaré, 70 (1999), pp. 473–495.

Downloaded 08/15/19 to 128.104.46.206. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.